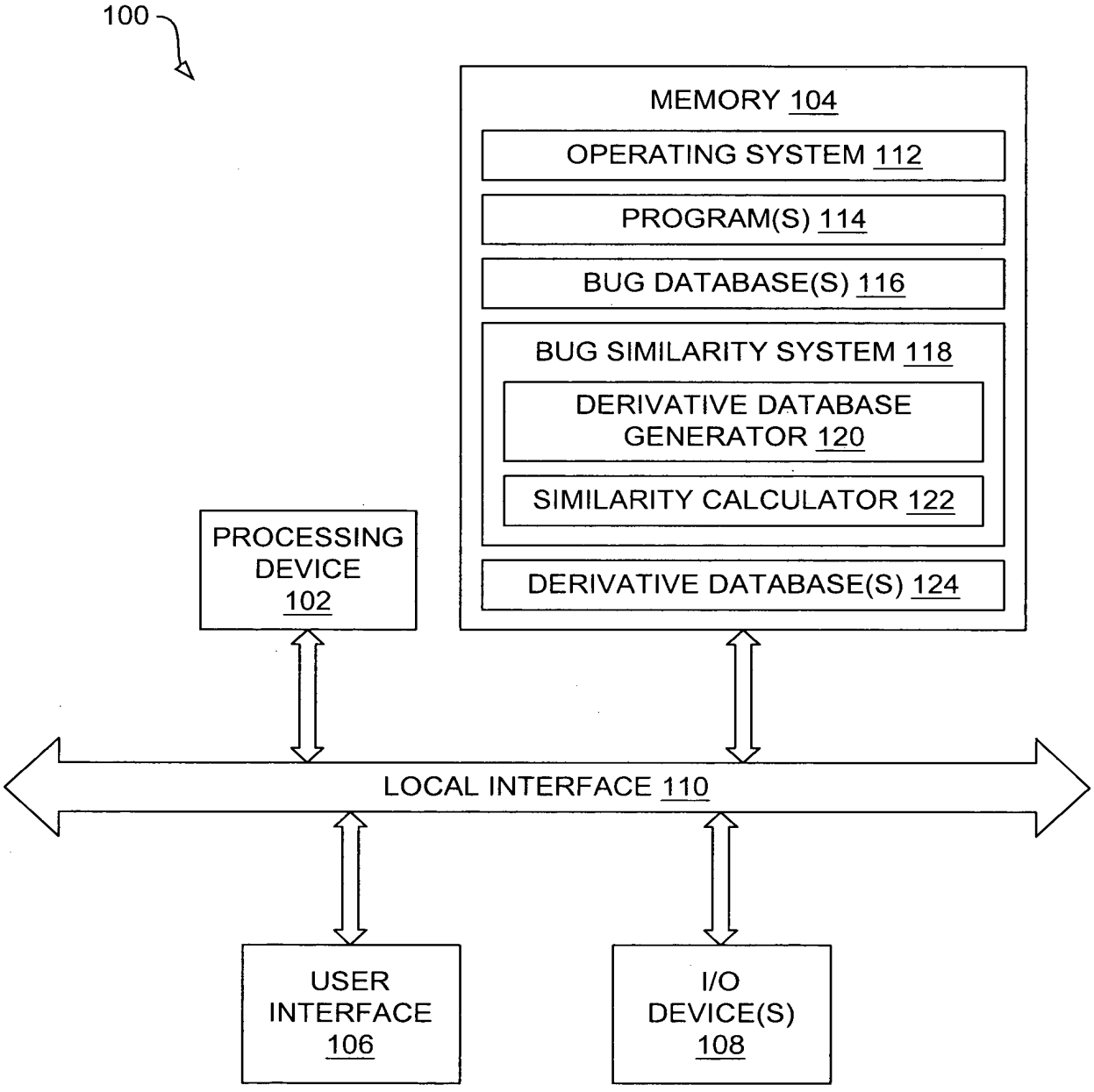
100
MEMORY 104
OPERATING SYSTEM 112
PROGRAM(S) 114
BUG DATABASE(S) 116
BUG SIMILARITY SYSTEM 118
DERIVATIVE DATABASE GENERATOR 120
SIMILARITY CALCULATOR 122
DERIVATIVE DATABASE(S) 124
PROCESSING DEVICE 102
LOCAL INTERFACE 110
USER INTERFACE 106
I/O DEVICE(S) 108

FIG. 1

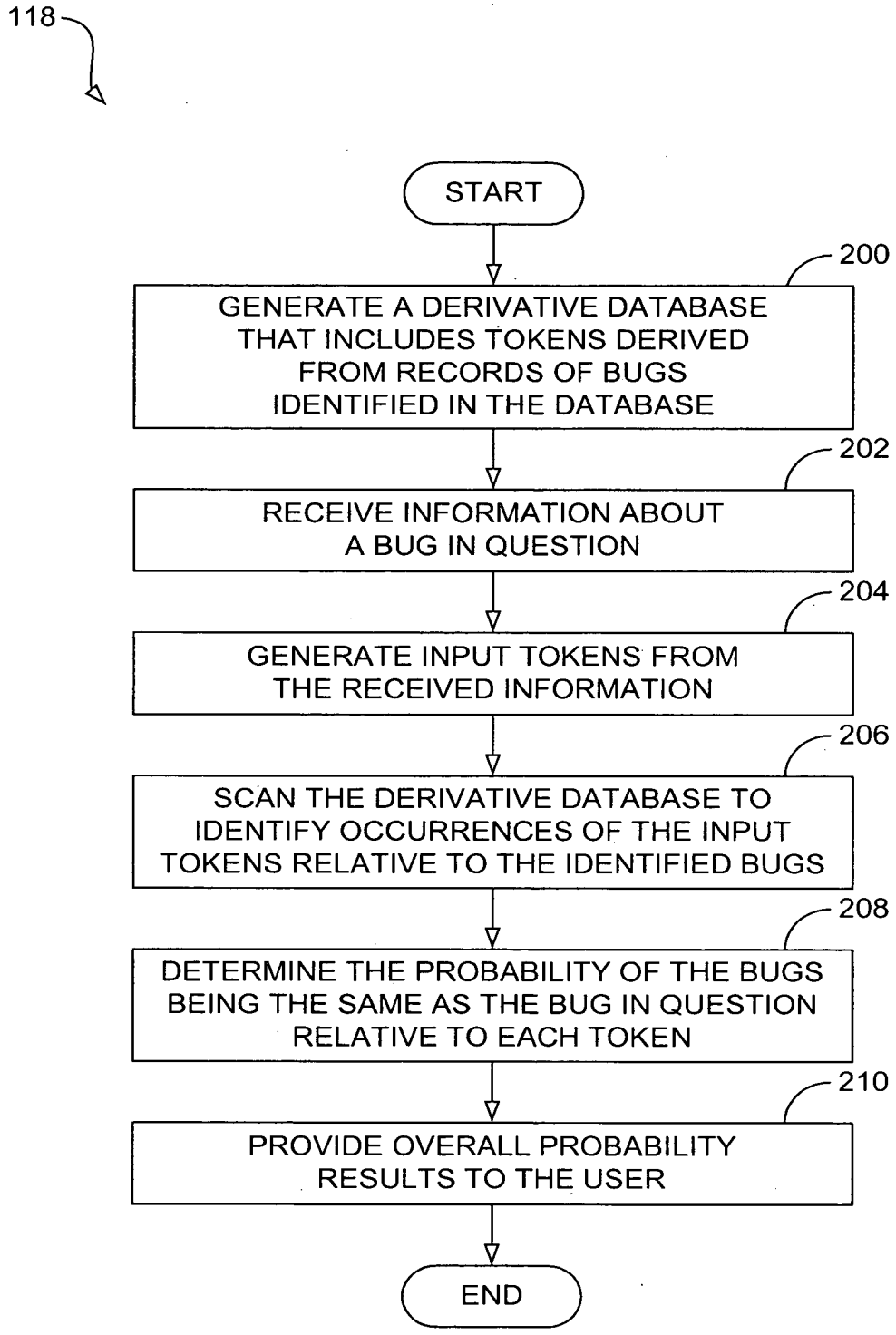
118
START
200
GENERATE A DERIVATIVE DATABASE THAT INCLUDES TOKENS DERIVED FROM RECORDS OF BUGS IDENTIFIED IN THE DATABASE
202
RECEIVE INFORMATION ABOUT A BUG IN QUESTION
204
GENERATE INPUT TOKENS FROM THE RECEIVED INFORMATION
206
SCAN THE DERIVATIVE DATABASE TO IDENTIFY OCCURRENCES OF THE INPUT TOKENS RELATIVE TO THE IDENTIFIED BUGS
208
DETERMINE THE PROBABILITY OF THE BUGS BEING THE SAME AS THE BUG IN QUESTION RELATIVE TO EACH TOKEN
210
PROVIDE OVERALL PROBABILITY RESULTS TO THE USER
END

FIG. 2

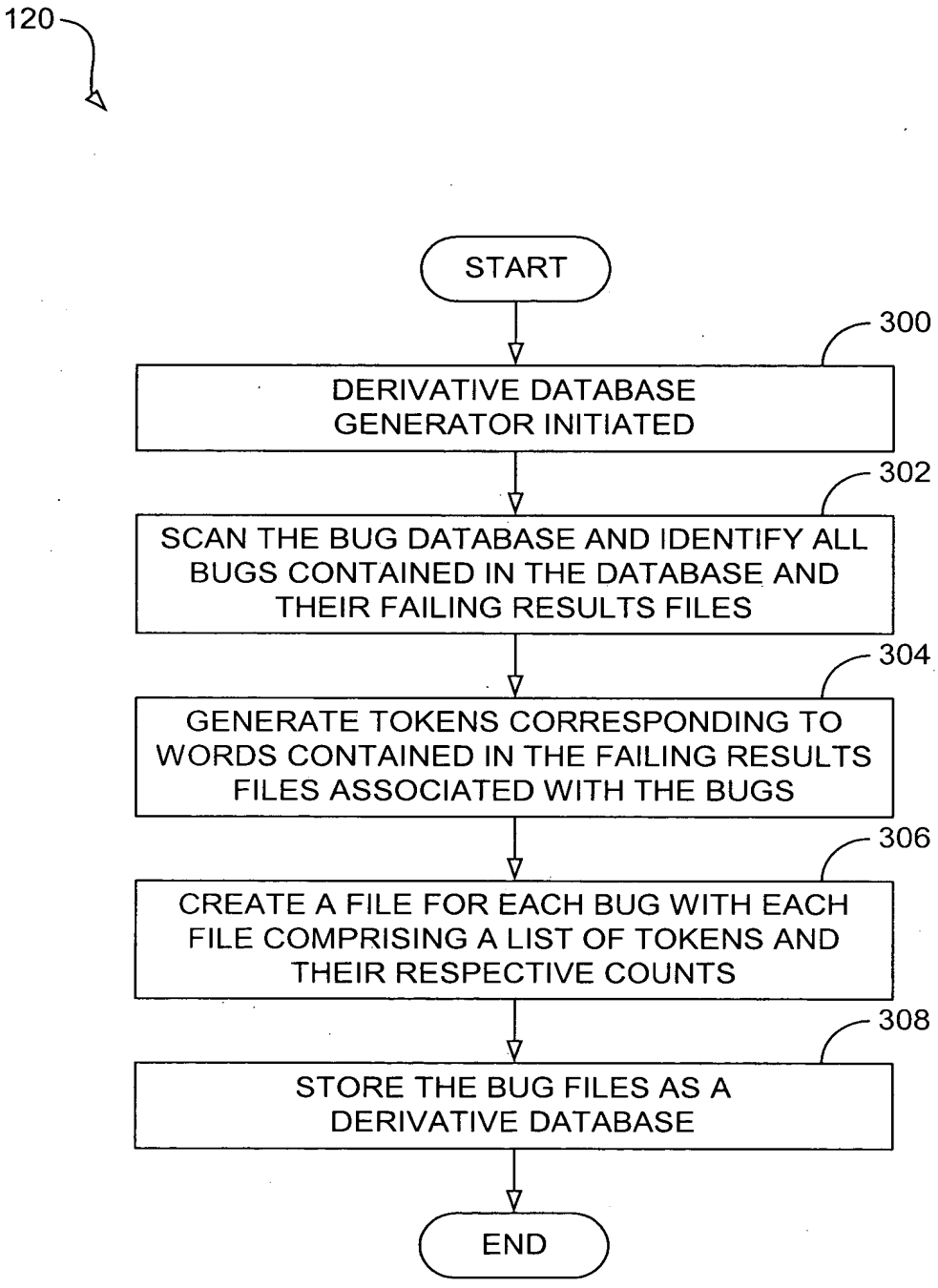
120
START
300
DERIVATIVE DATABASE GENERATOR INITIATED
302
SCAN THE BUG DATABASE AND IDENTIFY ALL BUGS CONTAINED IN THE DATABASE AND THEIR FAILING RESULTS FILES
304
GENERATE TOKENS CORRESPONDING TO WORDS CONTAINED IN THE FAILING RESULTS FILES ASSOCIATED WITH THE BUGS
306
CREATE A FILE FOR EACH BUG WITH EACH FILE COMPRISING A LIST OF TOKENS AND THEIR RESPECTIVE COUNTS
308
STORE THE BUG FILES AS A DERIVATIVE DATABASE
END

FIG. 3

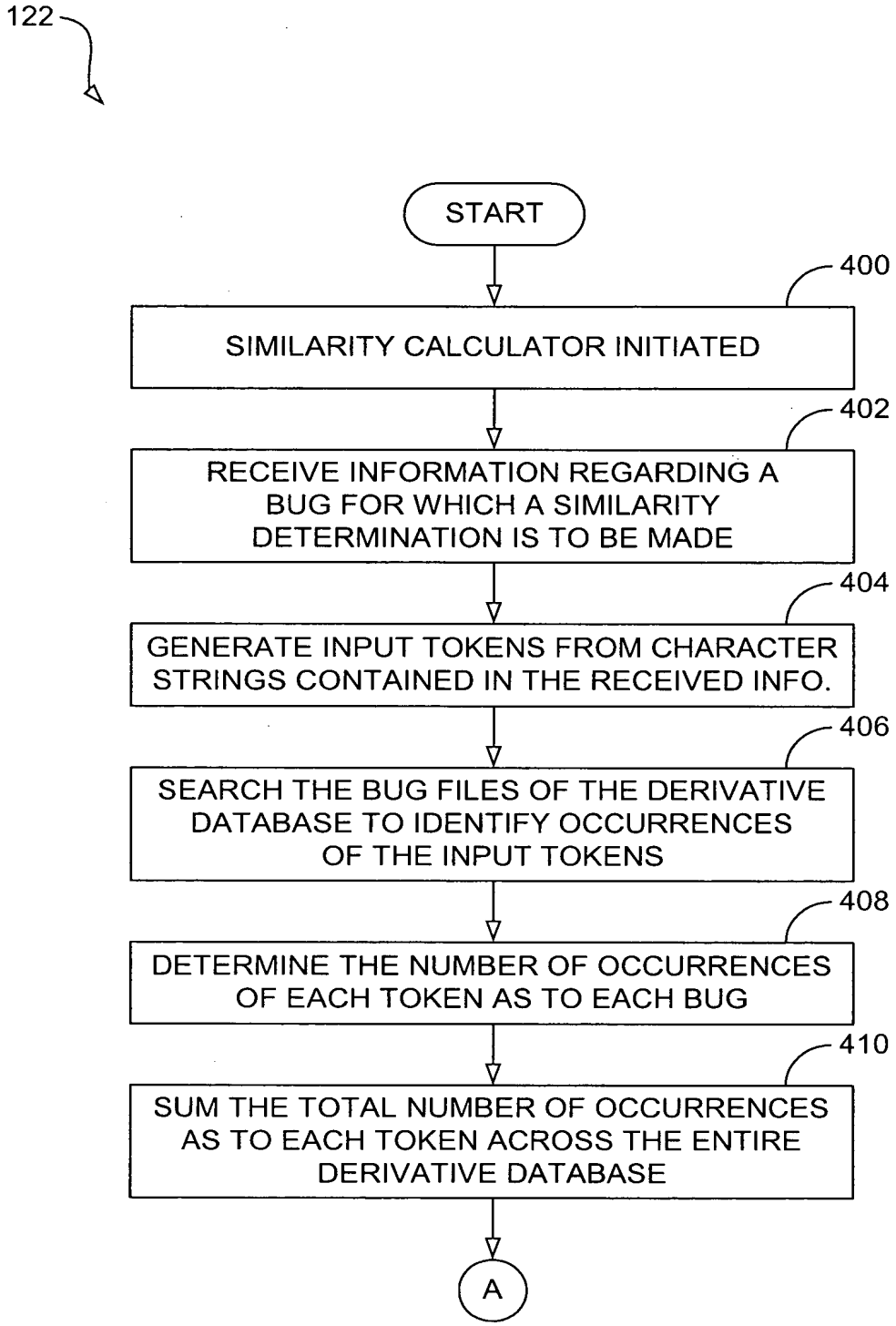
122
START
400
SIMILARITY CALCULATOR INITIATED
402
RECEIVE INFORMATION REGARDING A BUG FOR WHICH A SIMILARITY DETERMINATION IS TO BE MADE
404
GENERATE INPUT TOKENS FROM CHARACTER STRINGS CONTAINED IN THE RECEIVED INFO.
406
SEARCH THE BUG FILES OF THE DERIVATIVE DATABASE TO IDENTIFY OCCURRENCES OF THE INPUT TOKENS
408
DETERMINE THE NUMBER OF OCCURRENCES OF EACH TOKEN AS TO EACH BUG
410
SUM THE TOTAL NUMBER OF OCCURRENCES AS TO EACH TOKEN ACROSS THE ENTIRE DERIVATIVE DATABASE
A

FIG. 4A

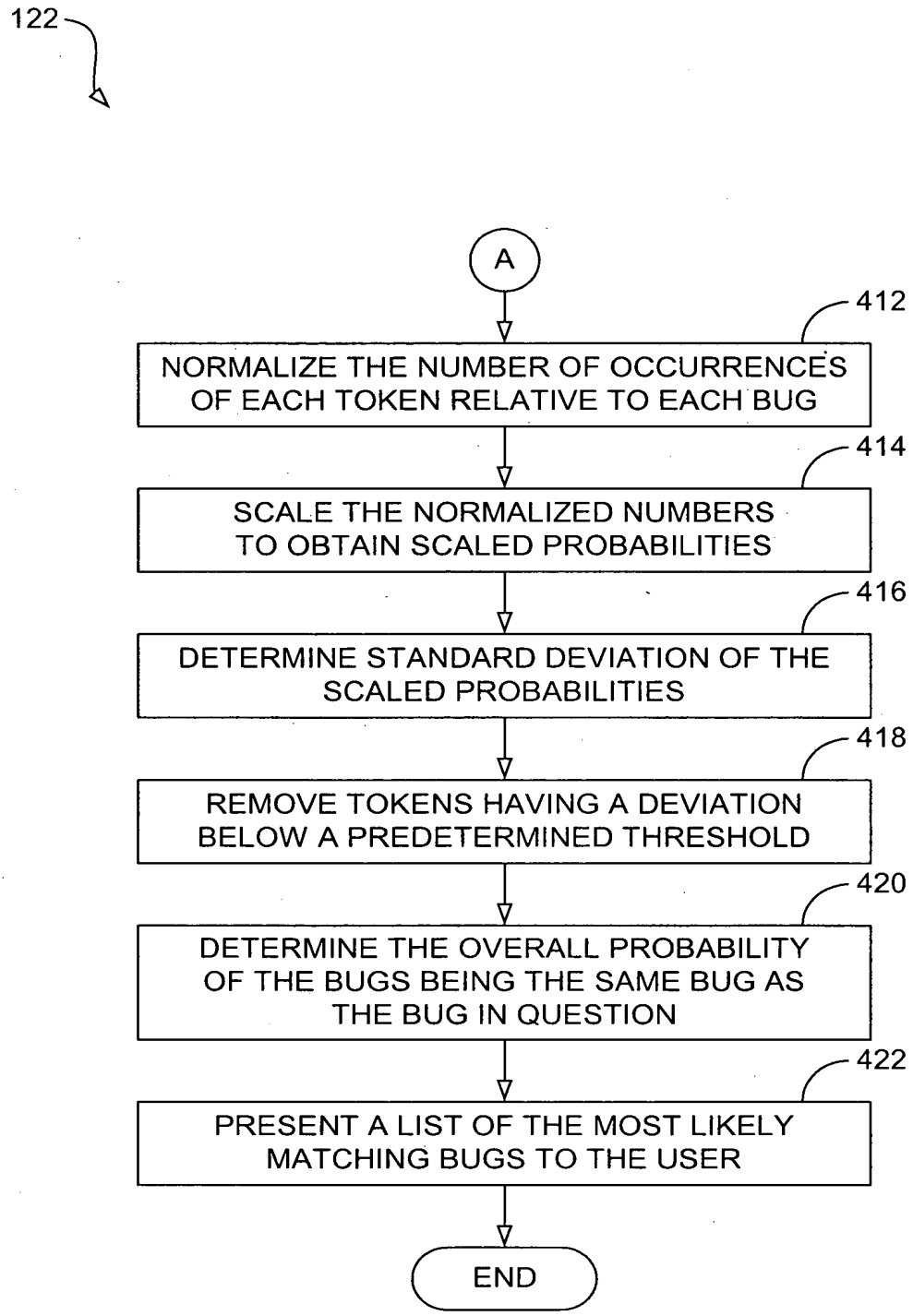
122
A
412
NORMALIZE THE NUMBER OF OCCURRENCES OF EACH TOKEN RELATIVE TO EACH BUG
414
SCALE THE NORMALIZED NUMBERS TO OBTAIN SCALED PROBABILITIES
416
DETERMINE STANDARD DEVIATION OF THE SCALED PROBABILITIES
418
REMOVE TOKENS HAVING A DEVIATION BELOW A PREDETERMINED THRESHOLD
420
DETERMINE THE OVERALL PROBABILITY OF THE BUGS BEING THE SAME BUG AS THE BUG IN QUESTION
422
PRESENT A LIST OF THE MOST LIKELY MATCHING BUGS TO THE USER
END

FIG. 4B

File	Token	Count
bug001	foo	5
bug001	bar	2
bug001	ack	7
bug002	foo	1
bug002	bar	2
bug002	ack	5
bug003	foo	4
bug003	bar	0
bug003	ack	3

FIG. 5A

Token	Owner	Count	Normalized Value
foo	bug001	5	0.50
foo	bug002	1	0.10
foo	bug003	4	0.40
bar	bug001	2	0.50
bar	bug002	2	0.50
bar	bug003	0	0.00
ack	bug001	7	0.47
ack	bug002	5	0.33
ack	bug003	3	0.20

FIG. 5B

Owner	Token	Scaled Probability
bug001	foo	0.50
bug001	bar	0.50
bug001	ack	0.47
bug002	foo	0.10
bug002	bar	0.50
bug002	ack	0.33
bug003	foo	0.40
bug003	bar	0.01
bug003	ack	0.20

FIG. 5C

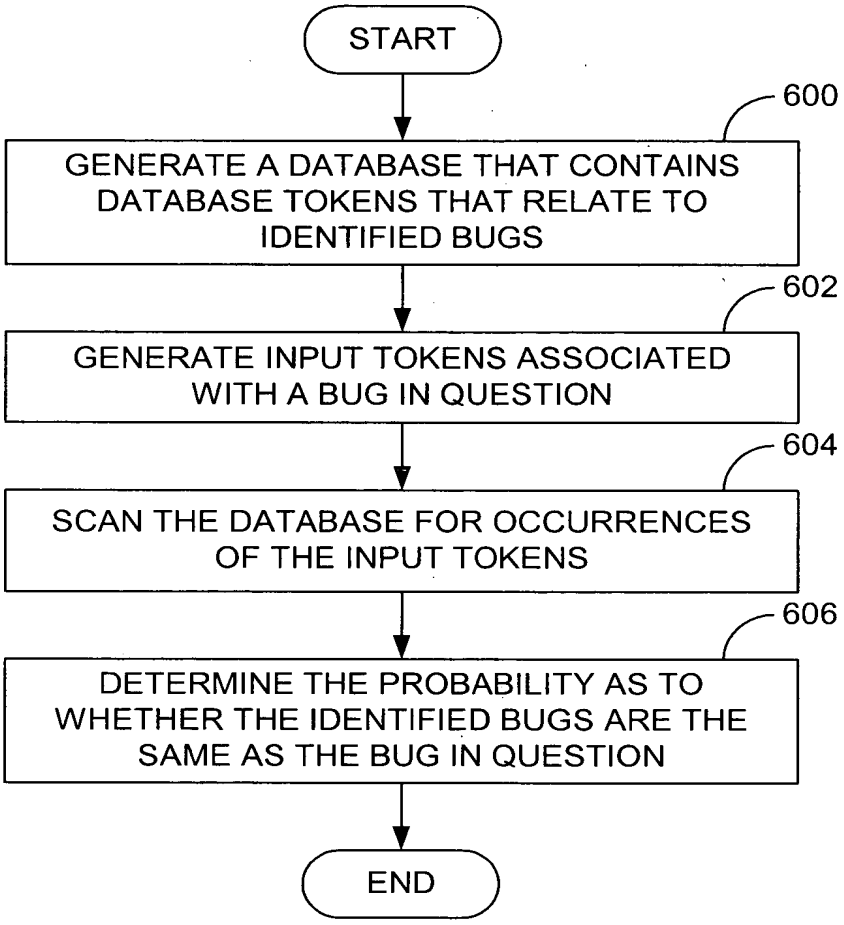
START
600
GENERATE A DATABASE THAT CONTAINS DATABASE TOKENS THAT RELATE TO IDENTIFIED BUGS
602
GENERATE INPUT TOKENS ASSOCIATED WITH A BUG IN QUESTION
604
SCAN THE DATABASE FOR OCCURRENCES OF THE INPUT TOKENS
606
DETERMINE THE PROBABILITY AS TO WHETHER THE IDENTIFIED BUGS ARE THE SAME AS THE BUG IN QUESTION
END

FIG. 6